

# Premonitory changes in seismicity prior to the Great Sumatra-Andaman earthquake of December 26, 2004

M. Imoto and N. Yamamoto

National Research Institute for Earth Science and Disaster Prevention, 3-1 Tennodai, Tsukuba-shi, Ibaraki-ken 306-0006, Japan

(Received May 18, 2007; Revised November 30, 2007; Accepted December 31, 2007; Online published July 4, 2008)

We studied changes in seismicity in a source region of the Sumatra earthquake prior to the earthquake on 26 December 2004. Rate increases were detected for earthquakes of  $M \geq 5.0$  using the ETAS method, with the assumption that independent events follow the Poisson process with at most only one rate change. Based on the declustered catalogue, we obtained consistent changes with simple Poisson models. Temporal changes in mean magnitude were obtained using the CUSUM method.  $b$ -values for three periods were calculated, taking into consideration the results of the CUSUM analysis, which showed a 10-year period of an anomaly high  $b$ -value followed by a 13-year period of a low  $b$ -value. The statistical significance of this change was confirmed by a bootstrap method. Both changes imply activation of seismicity in the source region.

**Key words:** Sumatra, premonitory change,  $b$ -value, ETAS, CUSUM.

## 1. Introduction

A giant earthquake occurred along the Sumatra-Andaman subduction on 26 December 2006 (Sumatra earthquake). This earthquake ruptured more than 1000 km of the eastern boundary of the Indian plate (Bilham, 2005; Lay *et al.*, 2005) and was the largest in terms of magnitude ( $M_w = 9.0$ – $9.3$ ) in the world since the 1964 Alaska earthquake. The mechanism of the earthquake suggests that the earthquake was an interplate event, resulting from the subduction of the Indian-Australian plate beneath that Andaman microplate. The aftershock area extended about 1200 km along the Andaman trough.

At the present time, it is widely accepted that we can only estimate the long-term probability of respective recurrences even when we have sufficient geological and historical earthquake data to estimate accurate recurrent intervals. In order to make the probability of an earthquake occurring useful in practice, such probabilities on shorter terms should be assessed through the premonitory stage of an earthquake. Recent improvements in the assessment of seismic hazards based on space and/or time patterns of regional seismicity may provide refinements of the probabilities on the shorter term.

Tanaka (2005) examined correlations of tidal phases with the timing of earthquakes that occurred before the Sumatra earthquake and found that the correlations are more significant as the time of the Sumatra earthquake approaches. His finding suggests that a critical level of stress appears over a wide area over the source region. Mignan *et al.* (2006) reported that the accelerated moment release (AMR) can be identified in the regions of the Sumatra subduction system

that must have been stressed before the Sumatra earthquake and that the effect of AMR is clearest for the epicentral regions at a level of statistical significance. If these observations indeed represent the true situation, other types of changes in seismicity should also be present in this region.

This paper focuses on possible changes in seismicity, such as rate changes, changes in  $b$ -values, and mean event size, which were associated with the Sumatra earthquake over a wide area of the seismic rupture zone. For conclusive results, it would be necessary to study both spatial and temporal variations of these parameters; given our limited sample, we are therefore constrained to a study of only temporal variations of the values. In this paper, we analyzed temporal changes in rate by two different methods: (1) the ETAS (Epidemic Type of Aftershocks) method for the original catalogue and (2) a simple Poisson model for a declustered catalogue. The results obtained from both methods are consistent. We also examine temporal changes in mean magnitude by the Cumulative SUM (CUSUM) method and calculate  $b$ -values for three periods divided by the suggested change points. We discuss the statistical significance of these changes with results of a bootstrap analysis. Thus, we were able to obtain reliable results of changes in rate and mean event size.

## 2. Data

For this analysis, we used a data set taken from the U.S. Geological Survey (USGS) National Earthquake Information Center, PDE Monthly Listing. We based our selection of aftershock hypocenters on the criteria that the earthquakes occurred (1) over a rectangular area (Fig. 1(a)), (2) at depths of 60 km from the surface (Araki *et al.*, 2006), and (3) between 01 January 1973 and 24 December 2004 (immediately prior to the Sumatra earthquake). The study area chosen was large enough to encompass the seismic precursory phenomena related to the Sumatra earthquake. This

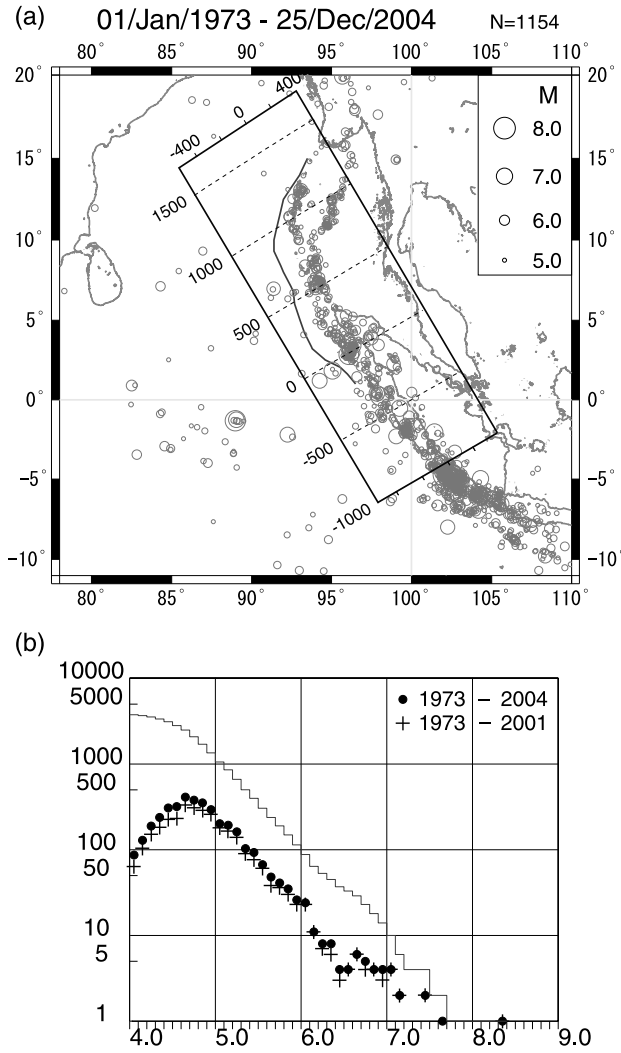


Fig. 1. (a) Epicenter map before the Sumatra earthquake along the Sumatra-Andaman trench. The box indicates the study area. (b) The Gutenberg-Richter magnitude frequency relation of earthquakes included in the study. The cumulative number curve is shown for earthquakes from 1973 to 2004. The magnitude-number plots for 1973–2004 is denoted by solid circles, and that for 1973–2001 is denoted by plus signs.

area included earthquakes governed by different kinds of stress regimes: the thrust type of the plate boundary, the ridge type in the back-arc opening, and the strike-slip type of the fore-arc slide. They were not analyzed separately in our study but used all together in the analysis since the data set did not contain a sufficient number of earthquakes for a statistical analysis of each group. Only events with a magnitude of  $\geq 5.0$  were used in this study. This cutoff magnitude was chosen by examining the magnitude frequency distribution of the catalogue in such a way as to observe a stable Gutenberg-Richter relation (Fig. 1(b)). For the shorter period of 1973–2001, a stable Gutenberg-Richter relation was still observed with this cutoff magnitude—with the exclusion of the event of high seismicity just before the Sumatra earthquake.

A catalogue declustered by the Reasenber method (1985) was also adopted, taking the effects of earthquake clustering into consideration. In a catalogue declustered by this method, the magnitude of a main shock is replaced with

that of equivalence to the total moment release in the respective cluster. However, we kept the original magnitude of the main shock for the following reasons. First, in an analysis of the Poisson process, we are not concerned with the magnitude of each event. Secondly, the formula used for calculating the moment strongly affects the value of the equivalent magnitude. The original formula employed in the method was obtained from the USGS Coda magnitude (Bakun, 1984), and it is clearly unsuitable in our case. Finally, in order to prevent the ambiguity of this point from influencing the results based on the declustered catalogue, we use the simplest approach, which is to maintain the original magnitude.

### 3. Changes in Rate

We assume that earthquake occurrences follow a stationary Poisson process; we denote the rate of this process as  $\mu_0$ . However, in practice, a clustering of earthquakes is often observed. Therefore, we further assume that each event may have its own daughter events as components of its magnitude. When we consider a change in rate, the Poisson rate,  $\mu_1$ , changes to another value,  $\mu_2$ , at a certain time. Considering clustering, we apply two different methods, the ETAS (Ogata, 1988) method and a declustered catalogue. Using the ETAS method, we can build clustering features into the analysis. If there is no change in rate, the hazard rate,  $\lambda_0$ , is given by

$$\lambda_0 = \mu_0 + \sum_{t_i < t} \frac{ke^{\alpha m_i}}{(t - t_i + c)^p} \quad (1)$$

where  $t_i$  and  $m_i$  are the time and magnitude, respectively, of the  $i$ -th event occurring prior to time  $t$  of assessment. In the case of a change, the hazard rate  $\lambda_1$  is given by

$$\lambda_1 = \begin{cases} \mu_1 + \sum_{t_i < t} \frac{ke^{\alpha m_i}}{(t - t_i + c)^p} & \text{for } t < t_c \\ \mu_2 + \sum_{t_i < t} \frac{ke^{\alpha m_i}}{(t - t_i + c)^p} & t_c \leq t \end{cases} \quad (2)$$

where  $t_c$  is the change point of the rate. Likelihood is represented by

$$L = \exp \left\{ - \int \lambda dv \right\} \cdot \prod_{i=1}^n \lambda(t_i), \quad (3)$$

where  $t_i$  is the time of the  $i$ -th event.

To select the better model, the Akaike Information Criterion (AIC; Akaike, 1977; Sakamoto *et al.*, 1983) was used. AIC is defined by

$$\text{AIC}_i = -2 \cdot \ln L_i + 2m, \quad i = 1, 2, \quad (4)$$

where  $m$  is the number of fitted parameters in the model and  $i$  refers to the first or the second model, respectively. In our case, there were five (six) free parameters for the first (second) model. The better model produces the smaller AIC. Therefore, when a difference in AIC between two models is defined by

$$\Delta \text{AIC} = \text{AIC}_1 - \text{AIC}_2, \quad (5)$$

a greater positive value of  $\Delta \text{AIC}$  suggests a change in rate. In our case, the optimal change point,  $t_c$ , is surveyed from a

Table 1. Difference in AIC between two models—change in rate and no change. ETAS indicates  $\Delta\text{AIC}$  values obtained by the ETAS analysis for the original catalogue. Declustered indicates  $\Delta\text{AIC}$  values obtained by simple Poisson models for the declustered catalogue. The change point,  $t_c$ , is indicated only for cases where  $\Delta\text{AIC}$  values are larger than 1.

Distance range	$M \geq 5$				$M \geq 5.5$			
	ETAS		Declustered		ETAS		Declustered	
	$\Delta\text{AIC}$	$t_c$	$\Delta\text{AIC}$	$t_c$	$\Delta\text{AIC}$	$t_c$	$\Delta\text{AIC}$	$t_c$
0–1500 km	2.30	26/Aug/2002	5.30	26/Aug/2002	3.00	5/Oct/2000	3.80	5/Oct/2000
–500–1500 km	1.57	25/Oct/2002	9.26	25/Oct/2002	3.00	9/Mar/2000	7.30	26/Aug/2002
–500–500 km	0.20	—	0.57	—	0.73	—	1.00	—
–1000–0 km	1.00	—	3.52	6/Aug/2000	0.26	—	0.83	—
–1000–1500 km	0.65	—	7.40	26/Aug/2002	0.57	—	4.06	26/Aug/2002

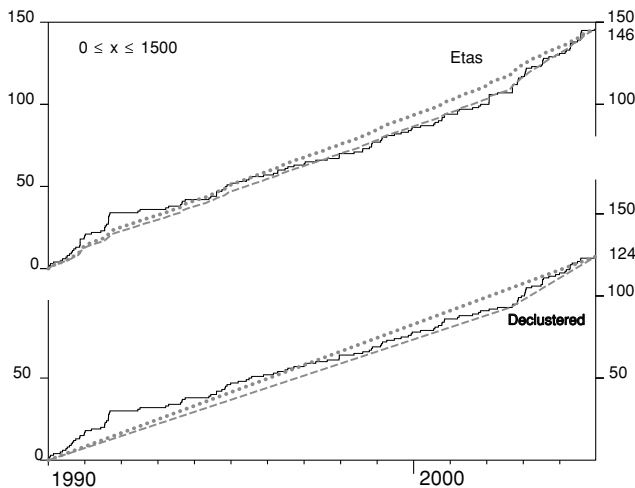


Fig. 2. Cumulative number curves for the region of 0–1500 km. The upper set depicts the cumulative number curve (solid) for the original catalogue and two theoretical curves (dashed and dotted lines), which are calculated by ETAS models with and without changes in rate. The lower set depicts the cumulative curve (solid) for the declustered catalogue and two theoretical curves (dashed and dotted lines), which are calculated by Poisson models with and without changes in rate, respectively. The change point of significance in rate was detected for both sets on 26 August 2002.

middle point of the study period (1 July 1997) to 30 September 2004 every 30 days in such a way that the maximum  $\Delta\text{AIC}$  is attained. When a declustered catalogue is analyzed, only the first term at the right side of Eq. (5) is considered.

Two different cutoff magnitudes, 5.0 and 5.5, and several different sampling zones along the N30W direction are adopted for both the original and declustered catalogues.  $\Delta\text{AIC}$  values are summarized in Table 1 and were obtained by the ETAS method from the original catalogue and the simple Poisson models for the declustered catalogue. Here, distance from the epicenter of the Sumatra earthquake is measured along the N30W direction.

The  $\Delta\text{AIC}$  values for two zones, 0–1500 km and –500–1500 km, respectively are large, and rate changes are statistically significant in the four different cases (two different cutoff magnitudes by two different methods). In other zones,  $\Delta\text{AIC}$  values in most cases are no larger than 1.0. This implies that the seismicity rates in those zones that correspond to the aftershock areas of the 26 December 2004 and the 28 March 2005 earthquakes became higher during

the last few years prior to the Sumatra earthquake. The optimal change points are determined in a period from March 2000 to October 2002, after which the rate increased by 60–100%. However, there were no rate changes in the extended zones of the aftershock area to the southeast, or, if they existed, they were not statistically significant.

Figure 2 illustrates the cumulative numbers of earthquakes at distances of 0–1500 km from the hypocenter of the Sumatra earthquake. The upper set depicts the observed cumulative curve (solid) and the two theoretical curves, change (dashed) and no change in rate (dotted), which were obtained by the ETAS method. The lower set depicts the cumulative curve (solid) and two theoretical lines (dashed and dotted), which were calculated by Poisson models with and without changes in rate for the declustered catalogue. The change point of significance in rate was detected for either set on 26 August 2002.

#### 4. Change in Mean Magnitude

The CUSUM method proposed by Page (1954) for the analysis of time series has been adopted in order to directly examine changes in mean magnitude. CUSUM can be defined for magnitude series  $M_i$ , as

$$C_i = \sum_{j=1}^i (M_j - M_{\text{ave}}), \quad i = 1, 2, \dots, n \quad (6)$$

where  $M_{\text{ave}}$  is the mean of the entire series.

Figure 3 illustrates the CUSUM for earthquakes ( $M \geq 5.0$ ) within 0–1500 km. The abscissa is the event number, and the vertical lines indicate the start of the year from 1975 to 2000 (every 5 years). Clearly, CUSUM indicates a decreasing trend from 1982 to 1990 and an increasing trend from 1990 to the time of the Sumatra earthquake. To examine the level of statistical significance, we generated 1000 series by a bootstrap method (Efron, 1979), each of which consists of 314 earthquakes identical to the original series. Each line of the figure connects the upper one percentile of  $C_i$ , the upper fifth percentile, the upper tenth percentile, the lower tenth percentile, the lower fifth percentile, and the lower one percentile for each event. The figure reveals that a major part of the CUSUM line after 1990 is lower than the lower tenth percentile line. This implies that the increasing trend after 1990 is highly significant.

Taking the CUSUM chart into consideration, we plotted  $b$ -values for three periods (corresponding to the three trends) in Fig. 4. Vertical lines indicate the uncertainty

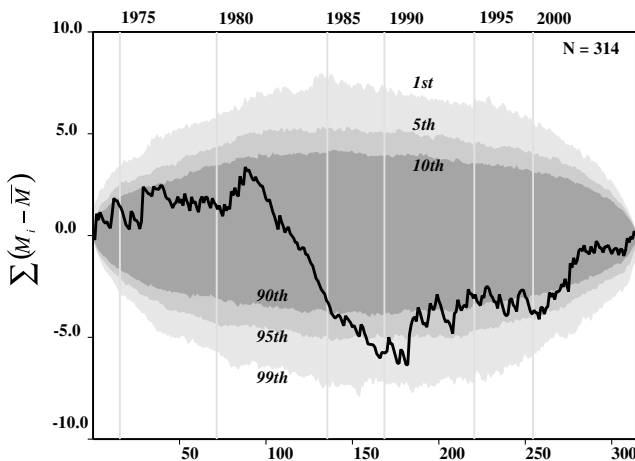


Fig. 3. CUSUM for earthquakes with a magnitude  $\geq 5.0$ . Each border line of the shadow part connects the upper one percentile of  $C_i$ , the upper fifth percentile, the upper tenth percentile, the lower tenth percentile, the lower fifth percentile, and the lower one percentile at each event. Each percentile is estimated from a bootstrap experiment with 1000 series of events.

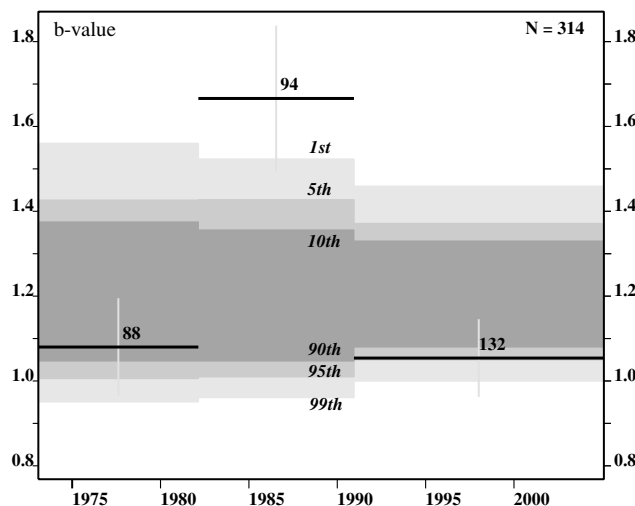


Fig. 4.  $b$ -values for three periods. The study period is divided into the times of the peak and the trough in the CUSUM chart in Fig. 3. The ordinate plots the maximum-likelihood  $b$ -value, with its standard error (Aki, 1965). Shadows indicate the upper one percentile of the  $b$ -value, the upper fifth percentile, the upper tenth percentile, the lower tenth percentile, the lower fifth percentile, and the lower one percentile.

range of each  $b$ -value. In the same way as in Fig. 3, the upper one percentile of the  $b$ -value, the upper fifth percentile, the upper tenth percentile, the lower tenth percentile, the lower fifth percentile, and the lower one percentile are shown. This plot indicates that a high  $b$ -value for the period of 1981–1992 is highly significant and that a relative low  $b$ -value for 1992–2004 is significant at the 5% level.

Figure 5 illustrates the CUSUM obtained from the declustered catalogue, which contains 250 earthquakes. The figure presents  $C_i$  values below the lower fifth percentile. The main features, decreasing and increasing trends, as in Fig. 3, are preserved in the figure. Figure 6 depicts three  $b$ -values (as in Fig. 4, together with six different percentiles). This plot confirms that a high  $b$ -value for the period 1982–1990 is statistically significant.

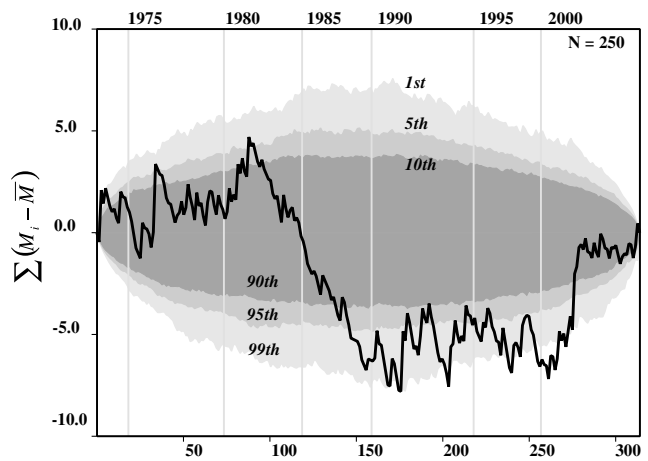


Fig. 5. Same as Fig. 3 but for the declustered catalogue.

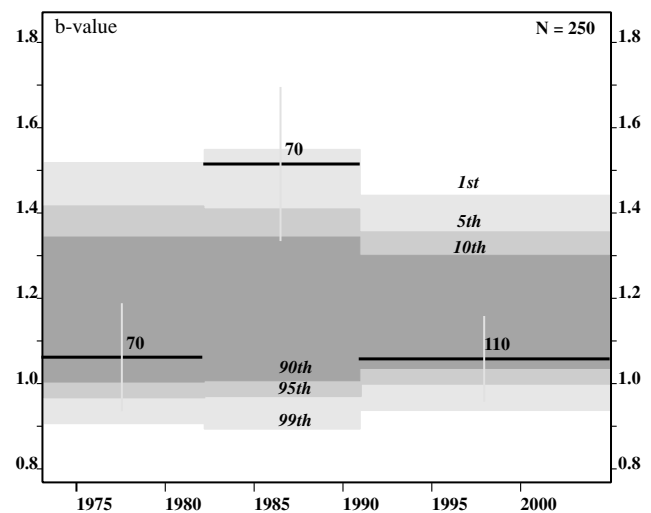


Fig. 6. Same as Fig. 4 but for the declustered catalogue.

## 5. Summary

The Sumatra earthquake was preceded by a period of an abnormally high  $b$ -value, which is statistically significant for both the original catalogue and the declustered catalogue. This finding may be consistent with the hypothesis that large earthquakes are preceded by periods of abnormally high  $b$ -values (Smith, 1998). This period may have begun as early as 1980 and continued until 1990, after which the mean magnitude increased to yet a higher value up until the time of the main earthquake.

This change in  $b$ -value implies that the seismic moment released by background seismicity accelerated from 1980 up to the time of the Sumatra earthquake—if seismicity rate in this period remained unchanged or increased with time. Actually, we detected seismicity rate increases 2–4 years prior to the Sumatra earthquake. Both findings are consistent with an AMR, which was reported by Mignan *et al.* (2006). These researchers studied AMR related to the Sumatra earthquake in detail with a statistical approach in approximate 250-km-long segments along the Sumatra–Andaman subduction and concluded that the AMR was clearest for the epicentral regions with a significance level

of less than 2%. We found changes in seismicity consistent with this conclusion using a popular and statistically comprehensive approach. Jiang and Wu (2005) concluded that a reliable preshock AMR process existed before this earthquake, with the duration of a quarter of a century. Their results suggest a premonitory change in seismicity of a much longer time duration than that indicated by the present results. This gap may have been due to the differences in study areas. Jiang and Wu (2005) used a circular area with radius of 800–1500 km centering at the hypocenter of the Sumatra earthquake. Their study area included earthquakes much further from the trench axis than those in our study, which were taken from the rectangular area 1000–2000 km along the trench. Seismicities far from the trench may transform the temporal variation into a gentler and longer range configuration.

In conclusion, we detected significant rate increases in the source region of the Sumatra earthquake 2–4 years prior to the event. We observed a 10-year period of an anomalous high  $b$ -value followed by a 13-year period of a low  $b$ -value. The statistical significance of this change was confirmed by a bootstrap method.

**Acknowledgments.** This manuscript was greatly improved by the comments of two anonymous reviewers and the editor Zhongliang Wu. This work was supported by the project of “The restoration program from giant earthquakes and tsunamis”, Special Coordinate Funds for Promoting Science and Technology, Ministry of Education, Culture, Sports, Science and Technology, Japan.

## References

- Akaike, H., On entropy maximization principle, in *Applications of Statistics*, edited by Krishnaiah, P. R., pp. 27–41, North Holland, Amsterdam, 1977.
- Aki, K., Maximum likelihood estimate of  $b$  in the formula  $\log N = a - bM$  and its confidence limits, *Bull. Earthq. Res. Inst. Univ. Tokyo*, **43**, 237–239, 1965.
- Araki, E., M. Shinohara, K. Obana, T. Yamada, Y. Kaneda, T. Kanazawa, and K. Suehiro, Aftershock distribution of the December 2004 Sumatra-Andaman earthquake from ocean bottom seismographic observation, *Earth Planets Space*, **58**, 113–119, 2006.
- Bakun, W. H., Seismic moments, local magnitudes, and coda-duration magnitudes for earthquakes in central California, *Bull. Seismol. Soc. Am.*, **74**, 439–485, 1984.
- Bilham, R., A flying start, then a slow slip, *Science*, **305**, 1126–1127, 2005.
- Efron, B., Bootstrap methods: Another look at the jackknife, *Ann. Stat.*, **7**, 1–26, 1979.
- Jiang, C. and Z. Wu, Test of the preshock accelerating moment release (AMR) in the case of the 26 December 2004 Mw 9.0 Indonesia earthquake, *Bull. Seismol. Soc. Am.*, **95**, 2016–2025, 2005.
- Lay, T., H. Kanamori, C. J. Ammon, M. Nettles, S. N. Ward, R. Aster, S. L. Beck, S. L. Bilek, M. R. Brudzinski, R. Butler, H. R. DeShon, G. Ekström, K. Satake, and S. Sipkin, The great Sumatra-Andaman earthquake of 26 December 2004, *Science*, **308**, 1127–1133, 2005.
- Mignan, A., G. King, D. Bowman, R. Lacassin, and R. Dmowska, Seismic activity in the Sumatra-Java region prior to the December 26, 2004 ( $M_w = 9.0$ – $9.3$ ) and March 28, 2005 ( $M_w = 8.7$ ) earthquakes, *Earth Planet. Sci. Lett.*, **244**, 639–654, 2006.
- Ogata, Y., Statistical models for earthquake occurrences and residual analysis for point processes, *J. Am. Stat. Assoc.*, **83**, 9–27, 1988.
- Page, E. S., Continuous inspection schemes, *Biometrika*, **41**, 100–114, 1954.
- Reasenber, P., Second-Order Moment of Central California Seismicity, 1969–1982, *J. Geophys. Res.*, **90**, 5479–5495, 1985.
- Sakamoto, Y., M. Ishiguro, and G. Kitagawa, *Akaike information criterion statistics*, 290 pp, D. Reidel, Dordrecht, 1983.
- Smith, W. D., Resolution and significance assessment of precursory changes in mean earthquake magnitudes, *Geophys. J. Int.*, **135**, 515–522, 1998.
- Tanaka, S., Tidal triggering of earthquakes precursory to the 2004 MW=9.0 off Sumatra earthquake, *AOGS 2<sup>nd</sup> Annual Meeting, Abstract*, 58-SE-A0453, 2006.

M. Imoto (e-mail: imoto@bosai.go.jp) and N. Yamamoto